Estimation of an Unknown Projection from a Map and its Applications

Tomas Bayer

Faculty of Science, Charles University in Prague, Czech Republic

Abstract. This paper presents new methods for the detection of an unknown map projection and its parameters from a map. Corresponding 0D-2D elements both on the analyzed map and a sphere (or a reference map) represent the source matter for the analysis. The following cartographic parameters (i.e., constants of a projection P) are estimated: R, φ_k, λ_k, φ_0, λ_0, Δx, Δy. Our solution minimizes L_2 norm of residuals and allows for the exclusion of incorrectly drawn elements from analysis. Both on-line and off-line methods of the detection are supported. Results are presented for early maps from the Map Collection of the Charles University and David Rumsay Map Collection. All algorithms were implemented in new detectproj SW, which supports more than 50 map projections and several operating systems.

Keywords: map projection, analysis, digital cartography, early maps, genetic algorithms, least square.

1. Introduction

Maps are an important part of our history and cultural heritage, there is great attention on their study and research. New methods and techniques for their analysis allow for the creation of full or partial geometric reconstruction of a map's content. This approach belongs to the category of cartometric analysis, whose capabilities with the rapid development of computer technology have been significantly increased.

The detection and estimation of an unknown cartographic projection and its parameters from a map represents a process of finding and establishing cartographic relationship between a map and the Earth. Such a type of analysis is beneficial and interesting for maps without any information about the used map projection. This applies particularly to historical maps, older maps or current maps. The aim of such an analysis is to determine a

cartographic projection used for a map construction and further improve its georeference.

For the georeferencing of maps covering a small territory (large- and mid-scale maps), the 1st order transformation is sufficient. Here an impact of the map projection can be neglected. However, this approach cannot be applied to small-scale maps (world maps, maps of continents or large countries) where a map projection influence should not be ignored.

Cataloging of early maps creates the need for additional cartographic information which is part of the meta data. In particular, they include information about a geographic extent, a map projection or a map scale. The bibliographic format Marc 21 contains a detailed description of a map projection in the fields 034 and 255B of the bibliographic record. Unfortunately, there was no method to determine these parameters in an accurate, fast, correct or for a large amount of maps. It is necessary to take into account that a cataloger can spend approximately 20 minutes with one map; therefore, the process of detection must be quick.

These requirements led to the development of new tools for on-line map projection analysis. Our paper does not describe the technical details, mathematical background nor the implementation specifics, which can be found in [**Fehler! Verweisquelle konnte nicht gefunden werden.**]. However, it familiarizes readers with examples, applications and practical outputs.

2. Related work

There are several software tools focused on georeferencing and cartometric analysis of old maps. MapRectifier (**Fehler! Verweisquelle konnte nicht gefunden werden.**) as well as WorldMap WARP (**Fehler! Verweisquelle konnte nicht gefunden werden.**) enable georeferencing of locally stored files. Several transformation models are supported; similarity, affine, spline. Some tools are part of more complex software packages, for example eHarta (**Fehler! Verweisquelle konnte nicht gefunden werden.**). Detection of an unknown map projection based on 2D transformations was used by (**Fehler! Verweisquelle konnte nicht gefunden werden.**). Algorithms have been implemented in the open source software MapAnalyst (**Fehler! Verweisquelle konnte nicht gefunden werden.**).

The Georeferencer (**Fehler! Verweisquelle konnte nicht gefunden werden.**), a tool based on the MapAnalyst engine, represents a new solution for on-line map analysis and collaborate georeferencing. However,

none of the presented SW supports neither transverse nor oblique aspects of a projection.

A mathematical background of the bellow described techniques can be found in **Fehler! Verweisquelle konnte nicht gefunden werden.**, **Fehler! Verweisquelle konnte nicht gefunden werden.**, **Fehler! Verweisquelle konnte nicht gefunden werden.**, **Fehler! Verweisquelle konnte nicht gefunden werden.**], **Fehler! Verweisquelle konnte nicht gefunden werden.**, **Fehler! Verweisquelle konnte nicht gefunden werden.**.

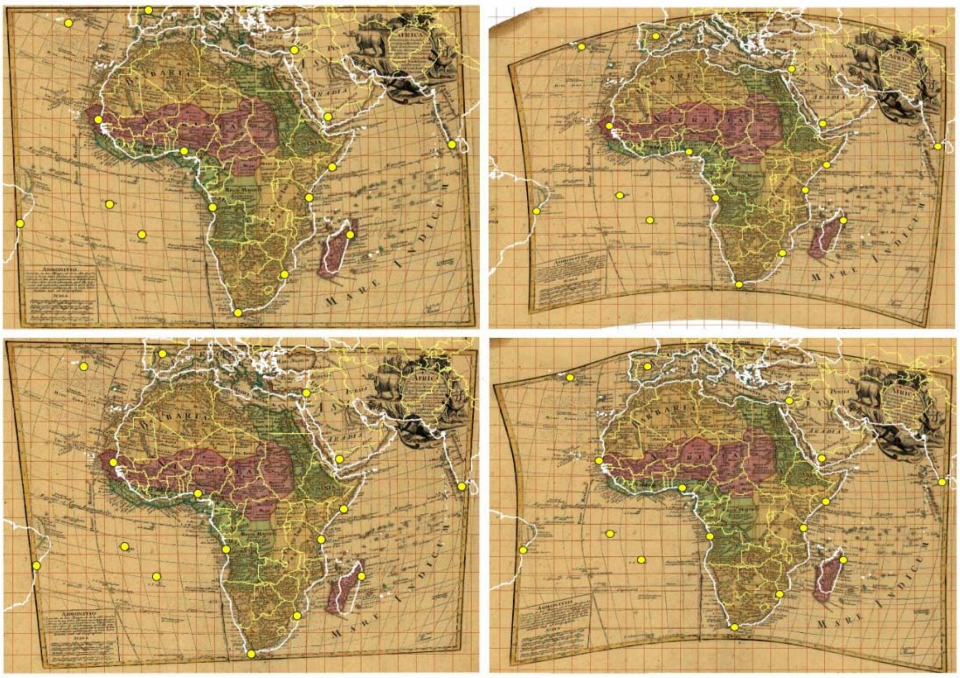


Figure 1. Georeferencing a map of Africa using various transformation models 1st, 2nd order similarity, projective and spline, reference map in Mercator projection.

3. Projection analysis and georeference

Before processing, a map needs to be georeferenced, where a correct geometric position in a coordinate system is established. The current approach, applicable to mid- or small-scale maps, is based on the application of several types of 2D transformations. To prevent geometric distortions of the map content, 1st order transformations (similarity, affine) are preferred. Here the influence of a map projection is to be neglected. For small territories

such a method is quite sufficient and appropriate. Finally, it has only a very limited application to maps showing small territories.

However, for small-scale maps, such an approach is completely inappropriate and wrong. Both analyzed and reference maps use heterogeneous coordinate systems, where no linear relationship between the systems exists. Higher-degree transformation may not be used for the analyses because of the unnatural distortions and twists of the map content. Let us take a closer look at **Figure 1**, where the early map "Africa Concinnata Secundum Observationes Membror...", Delisle Guillaume, from the Map Collection of the Charles University is to be georeferenced. Four transformation models: 1st, 2nd order similarity, projective and spline are applied and compared. The above mentioned disadvantages are clearly noticed; the map frame, meridians and parallels are twisted and distorted.

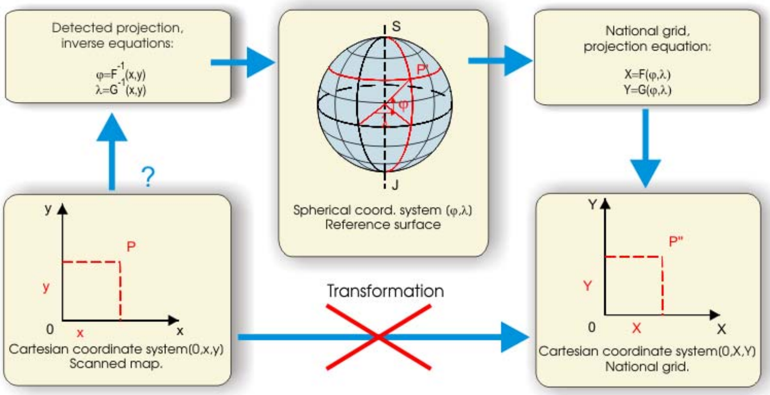


Figure 2. A correct georeference of early maps with determined projection parameters.

Proposed solution. The secondary deformation of the map content brings a geometric destruction of the map. To avoid this problem the following, more natural, solution is proposed, see **Figure 2**:

- Determine the analyzed map projection and its parameters of the map beeing gereferenced.
- Reproject the map to spherical coordinates φ, λ using inverse formulas (or re-project the current coordinate system to the map's projection).
- Reproject the map into the required coordinate system (a national grid).
- Correct additional shifts using the 1st order transformation.

Due to the difficulty of an unknown projection determination, which requires a deep numerical analysis, this problem has not been given attention so far. Finding unknown parameters represents a crucial point of the proposed procedure. **Figure 3** shows a re-projection of the current coordinate system to the map's projection. Here the estimated projection is Bonne applied in the normal aspect, where φ_k=90°N, λ_k=0.0°E, φ_0=25.6°N and λ_0=21.8°E. The result looks more natural than using a transformation.

Some parameters can be approximately found by an experienced cartographer. However, if a whole graticule is not available or a projection is applied in the oblique aspect; the correct values of the parameters are estimated by the trial and error method. Shapes of projected meridians, parallels or poles may make this process easier and help to exclude inappropriate candidates. In general, such an approach is tedious and desirable.

Therefore, a new method for on-line detection of projection parameters has been developed not dependent on the map scale, projection type and projection aspect robust to outliers.

Early maps and projections. The majority of early maps constructed in the 16th century do not have both solid geometric and geodesic bases. They represent more pictures and "art" then serious cartographic products. Here, it is impossible to think of the existence of a map projection. Although since 17th century maps have a graticule, the map content drawn without serious measurements is inaccurate. Unfortunately, most of the map projections from this period have only graphic or geometric descriptions. This fact is related mainly to the globular projections that are difficult to express by formal equations. They can be found in many world maps created by Jocodus Hondius, Georg Seutter or Guillame Delisle.

Figure 3. Re-projection of the current coordinate system to the estimated map's projection.

4. Analysis description

An essential step of the analysis is to find the proper geometric characteristics of elements both in the analyzed map P and the reference maps (or a sphere) Q to decide which map projection has been used or whether a map can be classified as unprojected. Analysis is invariant to the map scale, projection aspect or shifts and may be set as independent to the rotation.

Input features. Our solution takes into account the set of appropriate 0D-2D elements, preferably construction elements of a map (graticule) and the map content (rivers, roads, woods). The Cartesian coordinates [x,y] on the analyzed map and spherical coordinates [φ, λ] on the sphere (Earth) of the corresponding elements, are known. Involving line features into the assessment process reduces the discretization and significantly improves the results. Polygonal features allow one to analyze extensive parts of a map in a single step and represent the best matter.

It should be emphasized that lower efficiency was achieved, if analyzed features do not have good properties. A projection over a small territory up to $\Delta\varphi=\Delta\lambda=3°$, territory around the central meridian, prime meridian, equator, true parallel, north/south poles, meta center is hard to detect. Here, all map

projections have similar properties and the impact of a projection is below the graphical accuracy of a map.

4.1. Principle of analysis

A cost function f_c measures dissimilarity between P and $\mathbb{P}(Q)$, where $\mathbb{X} = \{R, \varphi_k, \lambda_k, \varphi_0, \lambda_0, \Delta x, \Delta y\}$ represents the vector of actually estimated parameters of a projection P. The aim is to find optimal values of parameters $\hat{\mathbb{X}}$ minimizing the cost function f_c

$$\hat{\mathbb{X}} = \arg\min_{\forall \mathbb{X}} f_c\big(P, \mathbb{P}(Q)\big).$$

For each analyzed projection $\mathbb{P}$ from the list of projections, a vector $\hat{\mathbb{X}}$ is determined. The vector $\bar{\mathbb{X}}$

$$\bar{\mathbb{X}} = \min_{\forall \mathbb{P}} f_c\big(\hat{\mathbb{X}}\big)$$

with lowest values of f_c relates to $\bar{\mathbb{P}}$, is assigned to the analyzed map.

As mentioned above, the cost function f_c takes into account the spatial distribution of 0D elements and shapes of 1D/2D elements. A suitable parameterization for 1D/2D elements based on the comparison of turning function $\theta(P_i)$, $\theta(P_i')$ of corresponding elements P_i, P_i', is used. Their similarity $d(P_i, P_i')$ is measured by

$$d(P_i, P_i') = \left(\int_0^1 |\theta(P_i)(s) - \theta(P_i')(s)|^2\, ds\right)^2.$$

Further details can be found in **Fehler! Verweisquelle konnte nicht gefunden werden.**, **Fehler! Verweisquelle konnte nicht gefunden werden.**. This descriptor is reliable and easy to compute. However, there are such situations, when this method does not improve results. Map projections belonging to the same category, cannot be successfully detected only by the graticule. Shapes of meridians and parallels are analogous; therefore, additional 1D/2D elements must be involved. For example, all meridians and parallels of cylindrical projections are represented by the lines.

Local vs. global minimum. Our analysis may be adapted to the problem of finding the local/global minimum of f_c. There are many approaches on how to solve this non-convex problem with or without explicit values of ∇f_c. The global optimizing method for off-line analysis is based on the genetic algorithm strategy (differential evolution). For on-line analysis the local optimizing strategy based on NLSP is used. Both methods are iterative.

Unlike other techniques, where only residuals between P and P' are measured, here the true spatial distribution of points is reflected. Spatial analyses are based on the parameters of Voronoi diagrams generated under P and P'. This strategy is more reliable and provides better results.

Heuristic approach. To speed-up the detection and exclude inappropriate values of determined parameters, a heuristic strategy is applied. There are several fast heuristic criteria; such as, the shape of meridians/parallels, matching ratio or standard deviation between P and $\mathbb{P}(Q)$ that decrease the computational speed.

The heuristic approach helps us to find the appropriate values of parameters in respect to cartographic habits and patterns related to local distortions. We want to avoid a pure geometric construct, which does not represent the cartographic rules. For these purposes, the variation criterion is used. Under analyzed territory divided into k pieces with mid points $P_i = [\varphi_i, \lambda_i]$ the global Airy criterion E is computed

$$E^2 = \frac{\sum_{i=1}^{k} \varepsilon_i^2}{k},$$

where

$$\varepsilon^2 = 0.5(|a-1| + |b-1|).$$

A projection is acceptable, if $E^2 < 1.0 \cdot 10^{-8} M$, where M represents a scale of a map. It is noticeable that the impact of the Airy criterion must be lower than the graphical accuracy of a map. Further technical details, math background, formula derivation and implementation specifics, can be found in **Fehler! Verweisquelle konnte nicht gefunden werden.**, **Fehler! Verweisquelle konnte nicht gefunden werden.**, **Fehler! Verweisquelle konnte nicht gefunden werden.**, **Fehler! Verweisquelle konnte nicht gefunden werden.**, **Fehler! Verweisquelle konnte nicht gefunden werden.**, **Fehler! Verweisquelle konnte nicht gefunden werden.**.

Outlier detection. Drawn elements on early maps constructed without solid geometric or geodesic basis may be influenced by errors. Unfortunately, this issue negatively affects the results of analysis. There is an effort to find and exclude blunders from the detection process. This problem can be transformed to outlier detection, where many different strategies have been developed. Based on the analysis of hundreds of early maps, a limit of errors was estimated at up to 20%. Here IRLS or M-estimators seem to be appropriate techniques. The modified Danish method was set as a primary tool for outlier detection. Weights of measurements suspected to be outliers are iteratively decreased, while weights of 'good' measurements are not changed. For detected outliers on the analyzed map, see **Figure 4**. Removing incorrectly drawn elements significantly refines the results.

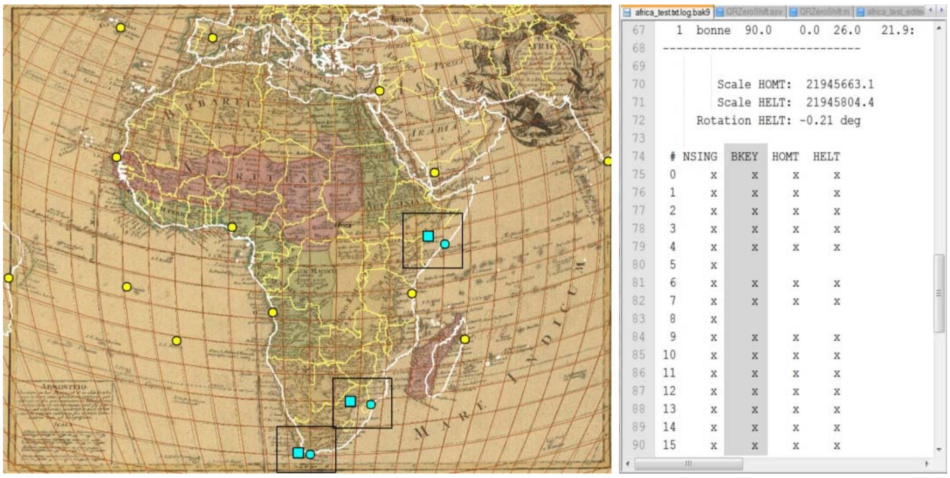


Figure 4. Detection of outliers on the analyzed map, 3 incorrectly drawn elements.

Recommendation for analysis. The efficiency of analysis depends on several factors, primarily both on the analyzed territory size and location. Small territories up to $\Delta\varphi = \Delta\lambda = 3°$ are undetectable as well as territories near the equator, central meridian or the north and south poles. Here an impact of a projection is lower than the graphical accuracy of the map. Analyzed features should be evenly distributed; the recommended amount of features is 10-15.

5. The software

The detectproj SW represents a new tool for the estimation of an unknown map projection and its parameters. It is based on the above mentioned analytical methods and supports both off-line and on-line detection strategies. The user interface is designed similar to the well-known Proj.4 library. It supports 55 map projections. On account of an unclosed solution, a definition of new map projections may be added. It supports several operating systems (Windows, GNU/Linux) and it is available for download free of charge.

Overview of the basic functions. There are several parameters and switches that allow for the configuration of input feature properties, detection method or heuristic sensitivity. Running and controlling the program is done from the command line

```
detecproj -switch +parameter=value test_file.txt  ref_file.txt
```

An input file with test points contains Cartesian coordinates [x,y] of all input points on the analyzed map. The x-axis always has east direction and

the y-axis always has north direction. Analogously, an input file with reference points contains spherical coordinates [φ,λ] of corresponding points. Let us briefly describe the most frequently used switches (see **Table 1**) and commands (see **Table 2**).

Switch	Description
-h	Enable heuristic, non-perspective samples are excluded from analysis
-n	Analysis in the normal aspect of a projection.
-t	Analysis in the transverse aspect of a projection.
-o	Analysis in the oblique aspect of a projection.
-r	Remove incorrectly drawn elements from analysis.

Table 1. The list of switches.

Both text and graphic results are provided. The output text file contains all relevant information about the detection process, a list of estimated map projections and their parameters. Values of members of the cost function, f_c, are sorted in ascending order by relevance. Graphical output is represented by a graticule generated over the analyzed territory. Both latitude and longitude increments can be specified by the user. Results are stored in DXF file and the overlay of the analyzed map and an estimated graticule in some CAD or GIS SW can be done. The graphic representation of results gives us a better overview and verification of determined parameters.

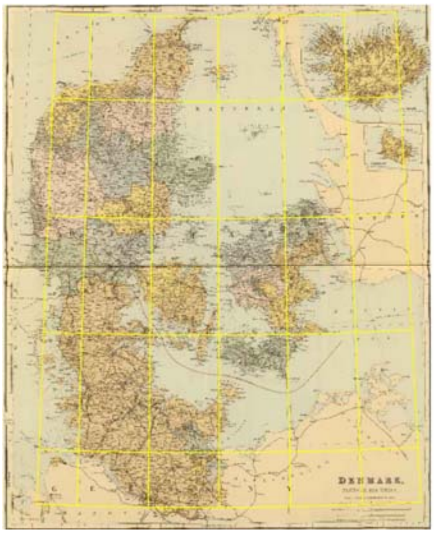
Parameter	Description
met	Select the method for analysis: m1 (NLSP) or m2 (GA).
res	Amount of printed samples
dlat	An increment of $\Delta\varphi$ between adjacent parallel in DXF file.
dlon	An increment of $\Delta\lambda$ between adjacent meridians in DXF file.
proj	Analyzed projection can be specified, name in accordance with Proj.4
latp	Latitude φ_k of the meta pole for analyzed projection can be specified.
lonp	Longitude λ_k of the meta pole for analyzed projection can be specified.
lat0	Latitude φ_0 of the true parallel for analyzed projection can be specified.
lon0	Longitude λ_0 of the central meridian for analyzed projection can be

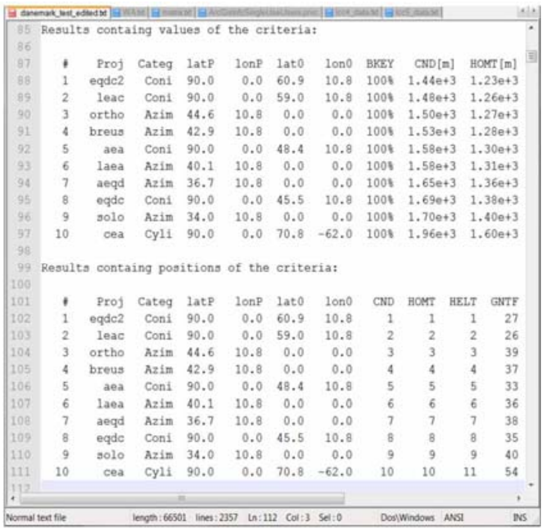
Table 2. The list of parameters.

Let us show an example, where a map of Denmark, in all aspects, is being analyzed. We want to determine the projection's parameters and draw a graticule with a step $\Delta\varphi=\Delta\lambda=1°$. The faster an on-line NLSP technique is chosen and no heuristic is applied. The command can be written as follows

```
detectproj danemark_t.txt danemark_r.txt -n -t -o +dlat=1 +dlon=1
```

Both text and graphic results are presented in **Figure 5**. The tables show sorted lists of the most probable results. The upper one contains estimated values of criteria and parameters, the lower one is sorted according to criteria. It is noticeable that there is a large consensus for the best sample, which had the best results in the most criteria. Only the turning function brought slightly discrepant values (see the last column). Thusly, all projection categories, projection names and parameters may be determined correctly. The geometric basis of this map is represented by the equidistant conical projection in the normal aspect, where the true parallel latitude is φ_o=61°N and the central parallel longitude is λ_o=11°E.





danemark_test_edited.txt

Results containg values of the criteria:

#	Proj	Categ	latP	lonP	lat0	lon0	BKEY	CND[m]	HOMT[m]
1	eqdc2	Coni	90.0	0.0	60.9	10.8	100%	1.44e+3	1.23e+3
2	leac	Coni	90.0	0.0	59.0	10.8	100%	1.48e+3	1.26e+3
3	ortho	Azim	44.6	10.8	0.0	0.0	100%	1.50e+3	1.27e+3
4	breus	Azim	42.9	10.8	0.0	0.0	100%	1.53e+3	1.28e+3
5	aea	Coni	90.0	0.0	48.4	10.8	100%	1.58e+3	1.30e+3
6	laea	Azim	40.1	10.8	0.0	0.0	100%	1.58e+3	1.31e+3
7	aeqd	Azim	36.7	10.8	0.0	0.0	100%	1.65e+3	1.36e+3
8	eqdc	Coni	90.0	0.0	45.5	10.8	100%	1.69e+3	1.38e+3
9	solo	Azim	34.0	10.8	0.0	0.0	100%	1.70e+3	1.40e+3
10	cea	Cyli	90.0	0.0	70.8	-62.0	100%	1.96e+3	1.60e+3

Results containg positions of the criteria:

#	Proj	Categ	latP	lonP	lat0	lon0	CND	HOMT	HELT	GNTF
1	eqdc2	Coni	90.0	0.0	60.9	10.8	1	1	1	27
2	leac	Coni	90.0	0.0	59.0	10.8	2	2	2	26
3	ortho	Azim	44.6	10.8	0.0	0.0	3	3	3	39
4	breus	Azim	42.9	10.8	0.0	0.0	4	4	4	37
5	aea	Coni	90.0	0.0	48.4	10.8	5	5	5	33
6	laea	Azim	40.1	10.8	0.0	0.0	6	6	6	36
7	aeqd	Azim	36.7	10.8	0.0	0.0	7	7	7	38
8	eqdc	Coni	90.0	0.0	45.5	10.8	8	8	8	35
9	solo	Azim	34.0	10.8	0.0	0.0	9	9	9	40
10	cea	Cyli	90.0	0.0	70.8	-62.0	10	10	11	54

Normal text file length : 66501 lines : 2357 Ln : 112 Col : 3 Sel : 0 Dos\Windows ANSI INS

Figure 5. Graphic and text results of analysis: generated graticule and text protocol.

6. Experiments and results

To demonstrate the capabilities of the software, three maps of different scales, sizes and projections have been used for tests. However, only the local minimizing NLPS strategy was involved. For all maps, the correct map projection parameters were not previously known. Analyzed maps belong to the Map Collection of the Charles University in Prague and David Rumsay Map Collection.

Map 1: "Europe Politique", Atlas St. Cyr. Furne, Jouvet et Cie, Paris, 1885. Estimated parameters of a projection: Bonne projection, φ_k=90.0°N, λ_k=0.0°E, φ_o=54.7°N, λ_o=20.2°N. The map has a geometric basis, the re-

sults are clear and the generated graticule fits to the analyzed map, see **Figure 6**.

Map 2: "Nova Totius Terrarum Orbis Geographica ac Hydrographica Tabula", Hendrik Hondius, 1630, Atlantis Maioris Appendix, Map Collection of the Charles University, east hemisphere. Estimated parameters of a projection: Stereographic projection, φ_k=-3.4°S, λ_k=56.7°E, φ_o=0.0°N, λ_o=0.0°E. The map does not have a solid geometric basis; rather, there is probably some kind of globular projection (detected as the stereographic projection very close to the transverse aspect). The absence of coordinate functions for such a projection causes the results to not be as clear. The generated graticule fits to the analyzed map slightly worse, see **Figure 7**.

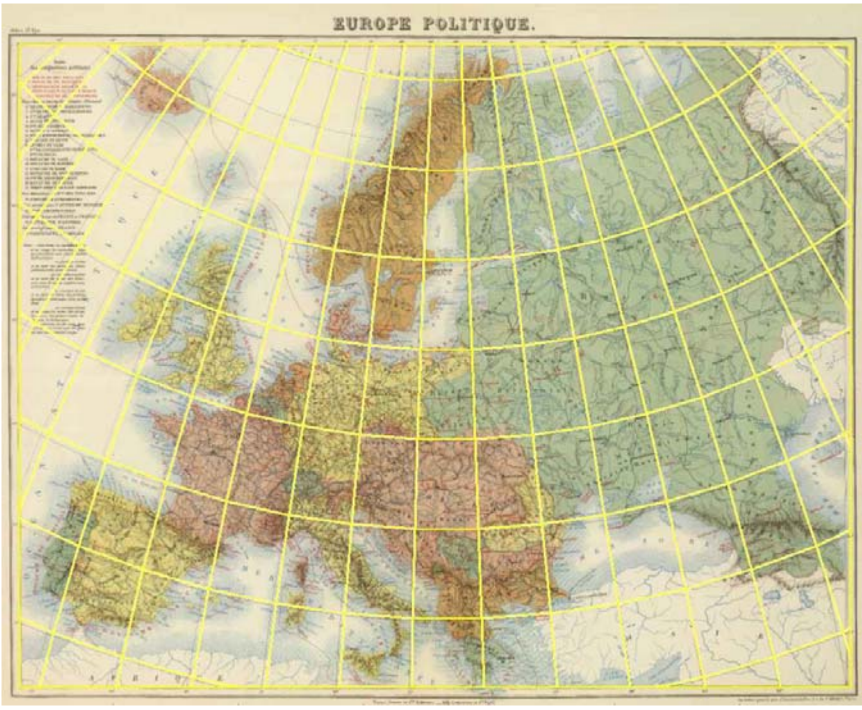


Figure 6. Generated graticule of Bonne projection (normal aspect) over the analyzed map.

Map 3: "British Islands", World Atlas, by A. Constable & Co. Edinburgh, 1817. Estimated parameters of a projection: orthographic projection, φ_k=42.3°N, λ_k=-2.7°W, φ_o=0.0°N, λ_o=0.0°E. The map has a solid geometric basis and the analyzed projection is in the oblique aspect. Generated graticule fits well to the analyzed, see **Figure 8**.

It is apparent that an on-line method of the detection based on the NLPS solution provides interesting results. The geometric reconstruction of pa-

rameters has a natural form. For maps with geometric basis there are no significant differences between actual and determined graticules (see maps 1,3). However, maps without geometric basis, as well as, maps using a graphical method of the projection (map 2), have some discrepancies between shapes of meridians and parallel. This problem applies particularly to globular projections, where due to the graphical construction, parametric equations are not available. But they can be accurately replaced by azimuthal projections in the transverse aspect.

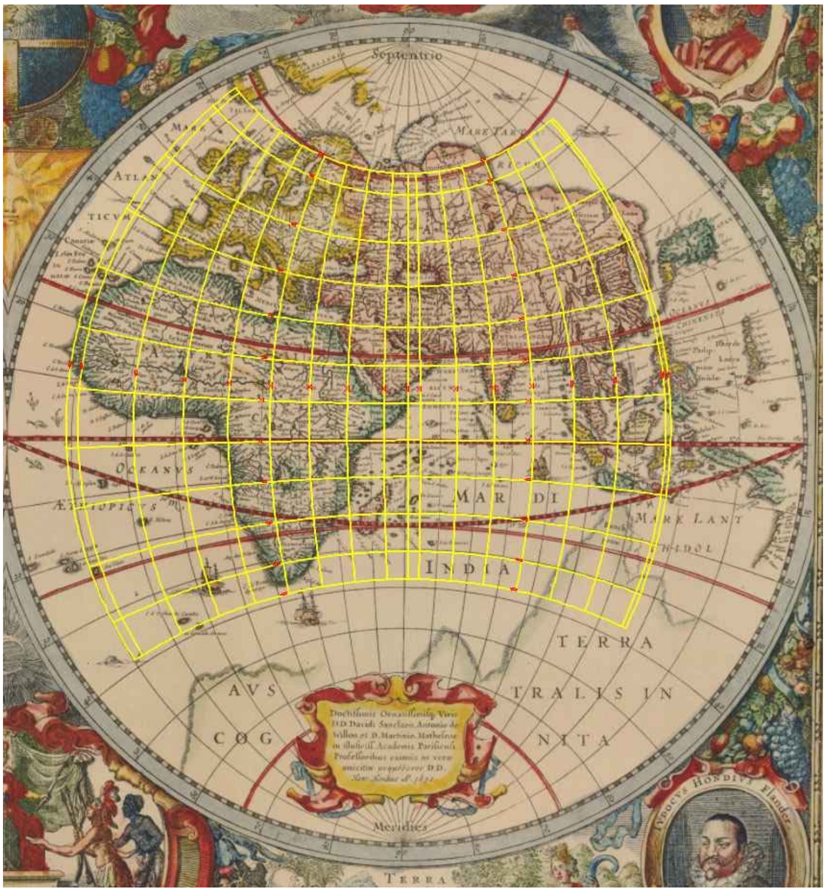


Figure 7. Generated graticule of the stereographic projection (close to the transverse aspect) over the analyzed map 2.

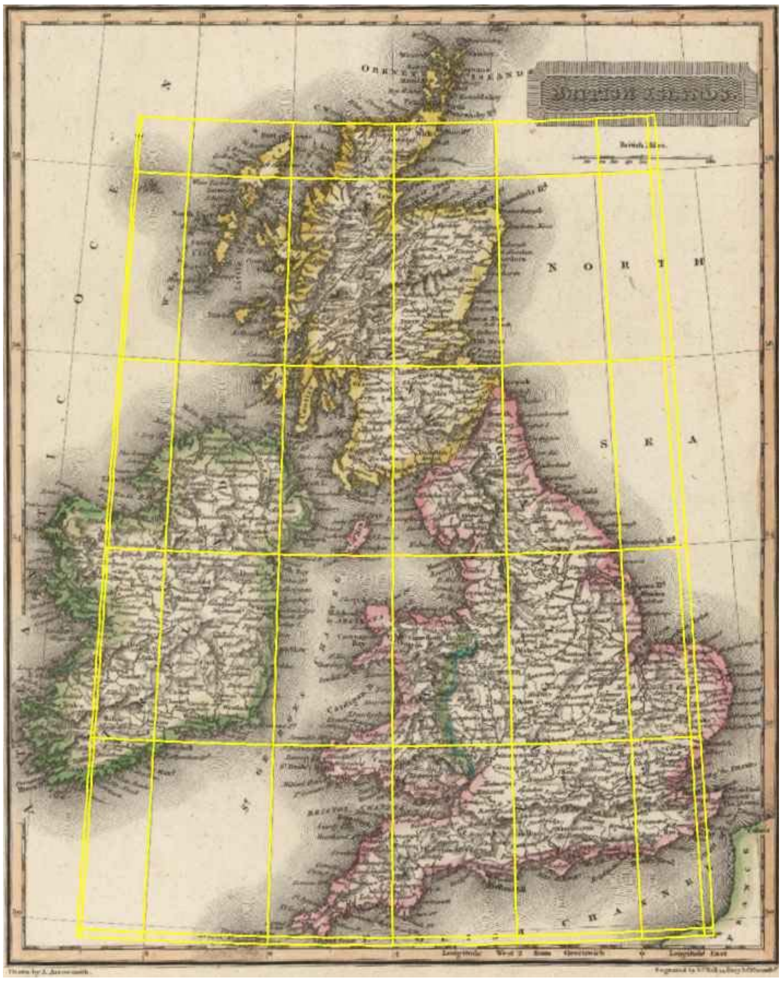


Figure 8. Generated graticule of orthographic projection (oblique aspect) over the analyzed map 3.

7. Conclusion

We briefly introduced a new method for an estimation of unknown cartographic projection parameters from a map. Our solution is based on robust statistical and numerical mathematics, and it provides both off-line and on-line methods of detections. The cost function f_c takes into account 0D-2D elements of the analyzed map. It does not represent a convex problem; moreover, it is poorly scaled and has large residuals. The on-line method based on NLPS strategy is to be stopped at some stationary point and it gives parameters of the local minimum. However, the off-line methods

based on the DE, found the global minimum of f_c, but it takes time. In most cases, the on-line method brings acceptable results and there are no significant cost differences between found global and local minimum (<0.05%). Our solution supports the elimination of incorrectly drawn elements from a map, which negatively affect the results.

It is important to point out that small territory up to $\Delta\varphi = \Delta\lambda = 3°$ is undetectable, as well as, a territory nearby the equator, central meridian or north/south poles, where most of the projections have similar properties.

Finally, neglecting a map projection cannot be applied to small scale maps (world maps, hemisphere maps, maps of continents or large countries), where the influence of a map projection cannot be ignored.

Both methods have been implemented in new detectproj software which is available from http://natur.cuni.cz/~bayertom/detectproj.html. The software is accessible for download free of charge.

8. Acknowledgement

This article was supported by a grant from the Ministry of Culture of the Czech Republic no. DF11P01OVV003 "TEMAP - Technology for access to Czech map collections: methodology and software for protection and re-use of national cartographic heritage".

9. References

Arkin, E. M., Chew, L. P., Huttenlocher, D. P., Kedem, K., and Mitchell J. S. B. (1991). An efficiently computable metric for comparing polygonal shapes. *13*(3), 209–216.

Bayer, T. (2013a). Advanced methods for the estimation of an unknown cartographic projection from the map. *Will be published in 2013*, 1–19.

Bayer, T. (2013b). Estimation of an unknown cartographic projection and its parameters from the map. *Geoinformatica, under review*, 1–22.

Crciunescu, V., Constantinescu, S. (2006). Eharta. http://earth.unibuc.ro/articole/eHarta?lang=en.

Erle, S., Krishnan, S., Waters, T. (2009). World map warp. http://warp.worldmap.harvard.edu/

Jenny, B. (2011). Mapanalyst- a digital tool for the analysis. http://mapanalyst.org/

Jenny, B., Hurni, L. (2011). Cultural heritage: Studying cartographic heritage: Analysis and visualization of geometric distortions. *Comput. Graph. 35*(2), 402–411.

Kelley, C. T. (1995). Iterative Methods for Linear and Nonlinear Equations. Number 16 in Frontiers in Applied Mathematics. SIAM.

Li, D., Fukushima M. (1999). A globally and superlinearly convergent gauss–newton-based bfgs method for symmetric nonlinear equations. *SIAM J. Numer. Anal. 37*(1), 152–172.

MetaCarta_Labs (2009). Map rectifier. http://labs.metacarta.com/rectifier/

Price, K. V. (1999). New ideas in optimization. Chapter An introduction to differential evolution, pp. 79–108. Maidenhead, UK, England: McGraw-Hill Ltd., UK.

Pridal, P. (2011). Georeferencer. http://www.georeferencer.org.

Qin, A. K., Huang, V. L., Suganthan, P. N. (2009, April). Differential evolution algorithm with strategy adaptation for global numerical optimization. *Trans. Evol. Comp 13*(2), 398–417.